DOCUMENT RESUME

ED 346 121 TM 018 372

AUTHOR LeCount, Judith; Fox, Paul W.
TITLE When Are Two Heads Worse than One? The Effect of Feedback and Structured Learning Groups on Confidence Calibrations.
PUB DATE Apr 92
NOTE 28p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 20-24, 1992).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Comparative Analysis; Confidence Testing; *Feedback; *Group Instruction; Higher Education; Individual Instruction; Instructional Effectiveness; *Knowledge Level; *Self Concept; *Undergraduate Students
IDENTIFIERS Accuracy; *Confidence Calibrations; *Structured Learning Groups

ABSTRACT

Students' confidence calibrations on quizzes taken individually and again in structured learning groups were examined. Under conditions of weekly feedback on accuracy and confidence, students working as individuals and in groups were hypothesized to improve not only their accuracy, but also their calibration of confidence. Findings were based on individual and group assessments of learning over a semester on actual course material. Data include confidence ratings from 32 undergraduates (13 males and 19 females) and their 10 learning groups on a series of 11 weekly quizzes. Within groups, students improved their accuracy, but they did not appropriately adjust their confidence judgments. The benefit of improved accuracy in groups came at the cost of increased confidence for wrong answers. Neither frequent feedback nor assignment to structured learning groups appears effective at improving students' calibration of confidence, and in particular cases, either may even make it worse. Five tables and seven figures present study findings. There is a 12-item list of references. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

JUDITH LECOUNT

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED346121

When Are Two Heads Worse Than One?

The Effect of Feedback and Structured Learning Groups on Confidence Calibrations

Judith LeCount and Paul W. Fox

University of Minnesota

Division J: Postsecondary Education - Paper Presentation

American Educational Research Association

San Francisco, California

April 22, 1992

Running head: CONFIDENCE CALIBRATION IN GROUPS

TM018372

BEST COPY AVAILABLE

Abstract

This investigation examines students' confidence calibration on quizzes taken individually and again in structured learning groups. Under conditions of weekly feedback on accuracy and confidence, students working as individuals and in groups are hypothesized to improve not only their accuracy, but particularly their calibration of confidence. Findings are based on individual and group assessments of learning over a semester on actual course material. Data include confidence ratings from undergraduates (N = 32) and their learning groups (N = 10) on a series of eleven weekly quizzes. Within groups, students improve their accuracy, but they do not appropriately adjust their confidence judgments. The benefit of improved accuracy in groups comes at the cost of increased confidence for wrong answers. Neither frequent feedback nor assignment to structured learning groups, appear effective at improving students' calibration of confidence and, in particular cases, may even make it worse.

When Are Two Heads Worse Than One?
The Effect of Feedback and Structured Learning Groups on Confidence Calibrations

At various times in our lives we are asked to predict the outcome of uncertain events. We make our intuitive estimates on the basis of information we think we know. Unfortunately, we frequently tend to be inappropriately overconfident when assessing the accuracy of our knowledge (see Lichtenstein & Fischhoff, 1977). This tendency toward overconfidence may be a particular problem in educational settings. For example, faculty may be poor judges when assessing their students' understanding of course related information (Fox & LeCount, 1991). In addition, students may be poor judges not only of their own understanding of course related general knowledge, but also of the knowledge their peers possess (LeCount & Fox, work in progress).

Though the phenomenon of inappropriate overconfidence in one's knowledge has been well documented in laboratory settings (e.g., Koriat, Lichtenstein, & Fischhoff, 1980), it has rarely been studied with actual classes and students in higher education. This is somewhat surprising because one of the major thrusts of the laboratory work on inappropriate overconfidence has involved speculation about study and learning problems poor metamemory skills may pose for the student learner. In one of the few studies of this problem in a classroom setting, Shaughnessy (1979) found a modest but statistically significant correlation between students' overall exam scores and their ability to discriminate appropriately between what they did and did not know. Unfortunately, Shaughnessy also observed a tendency for the difference between the confidence judgments students assigned to their correct and their wrong answers to *shrink* over the course (four mid-term exams taken about one month apart), rather than expand. Zechmeister, Rusch, & Markell (1983) tried a training procedure to improve students' appropriate assignment of confidence and hence decrease inappropriate overconfidence. In particular, the training program explicitly instructed students about the phenomenon of inappropriate overconfidence, and gave them practice with feedback regarding their confidence judgments at a special session. Unfortunately, the benefits of the training were modest, and in many respects the trained students

continued to assign confidence scores to their answers that differed little from those of the untrained group.

The purpose of the present inquiry is to improve students' calibration of confidence in their knowledge of what they are learning during an academic course of instruction. We define calibration of confidence in general as the ability of students to discriminate between what they know and what they do not know. The overall plan of our research allows for the comparison of students' individual and group judgments of confidence in their knowledge with their actual performance.

We examine two different ways to improve calibration of confidence. First, we offer students frequent, accurate, and informative feedback about their actual degree of understanding and the appropriateness of their confidence estimations concerning their answers to quiz questions. Importantly, this feedback is provided each week as the students progress though the academic term. Second, we place students into structured learning groups and encourage them to learn from one another in an effort to improve not only the accuracy of their answers, but particularly the calibration of their confidence judgments concerning those answers.

The present research draws its theoretical perspectives from two rather diverse fields of study, one concerning human metamemory and calibration of confidence and the other concerning group dynamics and decision making processes. The degree to which our predicted probability of being correct matches the actual probability of being correct is one definition of our "calibration of confidence" in our knowledge. As noted above, most people tend to be overconfident in their calibration, saying an item has a higher probability of being correct than is actually the case. This tendency for unwarranted confidence is thought to be caused in part by processes of inference and reconstruction during memory retrieval that introduce error (Fischhoff, Slovic, & Lichtenstein, 1977). Requiring individuals to list reasons why an answer may be wrong before giving their confidence rating tends to improve the calibration of confidence judgments (Koriat, Lichtenstein, & Fischhoff, 1980). Likewise, feelings of confidence in present knowledge influences the pursuit of new information and interpretations of that information (Fischhoff, Slovic, & Lichtenstein, 1977). Under conditions of frequent, informative, and weekly feedback, we hypothesize that participants

working as individuals and in learning groups should learn to adjust their confidence judgments about their answers so as to more accurately reflect the actual correctness of their decisions.

An understanding of group dynamics and how groups function is important for productivity and effective decision making. Within our educational system many postsecondary instructors are structuring their classroom learning environments with cooperative learning groups to improve students' decision making and learning processes (Cooper, et al., 1990; Johnson, Johnson, & Holubec, 1990). These structured learning groups allow students to work together to facilitate their own learning and the learning of their peers. In the present research, we assign students to structured learning groups in an effort to improve both their accuracy and, most particularly, their calibration of confidence.

We predict that accuracy for course material when students work in groups will be higher than when students work as individuals. Cohen (1986) reports that engaging students to talk and work together results in higher average gains on math tests that deal with concepts and problem solving. Engaging students in a group quiz should reinforce their learning of course concepts. In addition, we predict that group members will reinforce one another's correct answers and accurately solve wrong choices more frequently than they will select a wrong group answer.

Our predictions about students' confidence ratings are based on three manipulations: instructions, feedback, and experience. Instructions were given to participants encouraging them to mark higher confidence scores when their confidence in their answers was high and lower confidence scores when their confidence in their answers was the slightest bit uncertain. Written feedback about the appropriateness of their confidence ratings was given individually to every student each week. Experience with the quizzes, feedback, and instructions was acquired by the participants each week. We anticipate that the cumulative gains in general experience would cause participants to become correspondingly more skilled at assigning appropriate confidence judgments to their answers. We thus predict that under conditions of weekly personal feedback, instructions, and experience, individual participants (and their learning groups) will increase their confidence ratings when correct and decrease their confidence ratings when wrong.

When one or more individuals in a group has an answer that differs from other group members, this should signal some obvious uncertainty to the group. In this situation, the uncertainty within the group is concretely apparent. The groups should then mark lower confidence scores under conditions of visible uncertainty to a level even lower than participants working as individuals. We thus hypothesize that the use of groups within the classroom should help students adjust their confidence ratings to more effectively represent the true nature of their decisions. Figure 1 describes our predictions for individual and group confidence when answers are correct and wrong.

Insert Figure 1 about here.

Not all of our expectations were affirmed in this investigation, and in fact, some major surprises awaited us. However, the resulting information provides far more awareness about group processes and confidence calibration than we anticipated. We report what works well and what is unsuccessful for improving accuracy and confidence calibration in individuals working as individuals and within groups. We explore interesting disconfirmations to some of our hypotheses and we examine some intriguing issues in group dynamics for decision making processes in educational contexts.

Materials

Participants. The data for this investigation were collected during Spring Semester 1991 in an advanced educational psychology course composed of 33 St. Thomas University undergraduates of high academic standing. There were 20 women and 13 men, all of whom were accepted in a teacher preparation program for secondary education. One woman completed only half of the course and was dropped from this study (the final N=32).

Quizzes. Students received 1 quiz each week for a total of 11 quizzes beginning on week 2 and continuing through week 12. Each quiz consisted of 10 items of which 2 were true-false and 8 were multiple choice. Each item was followed by a confidence judgment scale that marked the limits of confidence in terms of the expected percentage of being correct about the answer (see Adams, 1957). For example, the confidence scale for the 2 true-false items began at 50%

(representing a completely chance level) and extended to 100% (indicating absolute certainty of correctness). The 8 multiple choice items had four alternatives for each item; therefore, the confidence scale began at 25% (representing a completely chance level) and extended to 100% (indicating absolute certainty of correctness). The students were instructed to mark at chance levels if they were completely unsure of an answer; they were told not to mark below chance at any time. Students were assured that marking their confidence levels in general and low confidence levels in particular would have absolutely no influence on their score in the class.

At the beginning of each class session, the quizzes were administered by the instructor along with verbal instructions for students to pay attention to their feelings of confidence about their answers and to depress their confidence score if they were the slightest bit unsure of their answer. An immediate reminder was written on the chalk board each week before the students took the quiz prompting them to mark their confidence scores as accurately as possible and to depress their confidence judgments if they felt the least bit uncertain in their answer.

Working as individuals. The students received a single answer sheet and quiz booklet. They were instructed to write their names on their answer sheet and quiz booklet. The students, as individuals, completed each quiz by marking their answers and rating their confidence that each answer was correct on their individual answer sheet. They were encouraged to mark their individual answers in their quiz booklets, but no further mention was made of their individual confidence ratings. The students were allowed 10 minutes (working individually) to complete each 10 item quiz. Upon completion of their individual quiz, one member of the group collected the individual answer sheets and returned these to the instructor. After the instructor gathered all of the individual answer sheets, the groups obtained a single, new answer sheet. All students retained their quiz booklets to take the same quiz again in their assigned learning groups.

Working in groups. Structured learning groups composed of 3 or 4 students were randomly assigned during the first week of class by the instructor. All groups were heterogeneous in terms of gender and race, but homogeneous in terms of age and experience. The groups were instructed to reach consensus on their individual answers and mark a group answer for each quiz item. They were also instructed to reach consensus on their group's confidence judgment for each item. The

groups were allowed 7 minutes to complete the group quiz and mark their group answers and confidence judgments on a single answer sheet. All members of the group participated in the group quiz and signed their names on the quiz answer sheet as an indication of individual participation in their group's work. One member of the group collected the group answer sheet and the individual quiz booklets and returned these items to the instructor. If the groups scored 80% or more on the group quiz, one bonus point was added to each individual group members' individual quiz score. Bonus points were frequently earned by all of the groups on each quiz. Students were assured that marking low group confidence levels would in no way influence their grade in the class. They were encouraged to mark their group confidence scores as accurately as possible and to depress their confidence judgments if they felt the slightest bit uncertain in their answer.

Feedback. Feedback on confidence and accuracy was both formal (i.e., from the instructor to the students one week later, item by item, on the previous week's quiz) and informal (i.e., within their structured learning groups). Students received formal feedback not only about the accuracy of each decision in terms of the number of items answered correctly, but also concerning the adequacy and appropriateness of their confidence judgments both as an individual and as a group member. When the confidence judgments were in the correct direction (high confidence when correct or low confidence when wrong), comments were written at the top of their individual and group answer sheets (e.g., "Your confidence scores are in the right direction"). When the confidence judgments were in the wrong direction (high confidence when the answer was wrong), comments were written at the top of their individual and group answer sheets encouraging students to suppress their confidence scores under conditions of uncertainty (e.g., "Your confidence judgments are in the wrong direction"; "Depress your group confidence score if you are the slightest bit uncertain in your response"; etc.). When mean confidence judgments for correct and wrong items were close together (span of 10% or less), comments were written at the top of their individual and group answer sheets encouraging students to spread their confidence scores apart by marking higher confidence scores under conditions of certainty and lower confidence scores under conditions of uncertainty (e.g., "Your confidence judgments are too close together"; "Depress your confidence score if your group is uncertain"; etc.). When mean confidence judgments for correct and wrong

items had a span of 20% or more, positive comments encouraging students to continue the rating of higher confidence scores when certain and lower confidence scores when uncertain were written at the top of their individual and group answer sheets (e.g., "Good confidence discrimination").

The individual and group answer sheets were hand scored by the instructor during the week and returned to the students after they completed their current week's individual and group quizzes. Thus, feedback on individual and group accuracy and confidence ratings for the immediate quiz has a delay of 2 quizzes. For example, Quiz 2 was completed prior to receiving feedback on Quiz 1; thus, feedback on Quiz 1 accuracy and confidence ratings is available the week before students take Quiz 3.

After completion of the sixth quiz, the groups were restructured to allow students to work with other members of the class. These restructured groups remained together to complete the second half of the course. This second set of groups was randomly assigned by the instructor based on the condition that students had had no past history of group work together in the course. Assignment of members to their new groups was made from randomly selected names on randomly selected group answer sheets. These new groups were heterogeneous on gender and race.

Results

Accuracy. The students working in groups consistently obtained higher mean accuracy scores than students working as individuals. Figure 2 shows that placing students in structured learning groups improved the accuracy of answers on every quiz. The mean quiz accuracy over all 11

Insert Figure 2 about here.

---

quizzes for participants working in groups significantly exceeded the mean quiz accuracy for participants working as individuals ($t_{(10)} = 6.902$, $p < 0.001$). Table 1 presents the mean quiz accuracy for individuals and groups overall.

Insert Table 1 about here.

The quizzes proved to be highly consistent in terms of item difficulty over the 11 week period. Using Quiz 6 as our pivot, we tested the first set of quizzes ($Time1_{(Quizzes\ 1\text{-}5)}$) against the last set of quizzes ($Time2_{(Quizzes\ 7\text{-}11)}$). Table 2 shows that during Time1 and Time2 there were no statistically significant differences in quiz difficulty for the 32 individual students ($t_{(31)} = .21$, $p = 0.839$). Likewise, there were no significant differences between the first set of 5 quizzes taken by the groups and the second set of 5 quizzes taken by the groups ($t_{(8)} = 0.202$, $p = 0.805$).

Insert Table 2 about here.

Confidence. We classified confidence judgments in terms of 3 categories: (1) overall confidence, (2) confidence when students or groups answered quiz items correctly, and (3) confidence associated with quiz answers that were wrong. Group mean confidence exceeds the individual mean confidence in each category. Mean confidence judgments for students working as individuals and in their groups are presented in Table 3.

Insert Table 3 about here.

Overall Confidence. Confidence judgments associated with all items determine an overall confidence whether they are answered correctly or incorrectly. Students working in groups achieved mean overall confidence scores significantly higher than those of students working as individuals ($t_{(10)} = 14.932$, $p < 0.001$). On every quiz the mean overall confidence for the participants working in groups exceeded the individual mean overall confidence. Thus, placing students in structured learning groups increases overall confidence on quiz responses. Although group overall confidence increased slightly from Time1 to Time2 (as shown in the lower panel of Table 4), this increase was *not* significant ($t_{(8)} = -1.445$, $p = 0.187$). However, there was a significant increase in overall confidence for the 32 individual participants from Quizzes 1-5 to Quizzes 7-11 (top panel of Table 4, $t_{(31)} = -2.32$, $p = 0.027$).

Insert Table 4 about here.

Confidence When Correct. As Tables 3 and 4 show, confidence when correct was high for participants working as individuals and even higher for participants working in their groups ($t_{(10)} = 12.178$, $p < 0.001$). Moreover, as shown in Figure 3, the groups exceeded the individual mean confidence when correct for each of the 11 quizzes. Achieving high confidence when correct is what we predicted would happen for individuals and their groups, although the students began at a much higher level of confidence far earlier in the study than we anticipated. Individual mean confidence when correct over $Time1_{(Quizzes1\text{-}5)}$ and $Time2_{(Quizzes7\text{-}11)}$ does increase, but this increase is shy of statistical significance ($t_{(31)} = -1.78$, $p = 0.085$). Group confidence when correct was high and not significantly different from $Time1_{(Quizzes1\text{-}5)}$ to $Time2_{(Quizzes7\text{-}11)}$ ($t_{(8)} = -1.059$, $p = 0.321$).

Insert Figure 3 about here.

Confidence When Wrong. Surprisingly, confidence when wrong increased during the quizzes for both the individuals and their groups. Moreover, as presented in Figure 4, group mean confidence when wrong exceeded individual mean confidence when wrong on 10 of the 11 quizzes.

Insert Figure 4 about here.

Furthermore, group mean confidence when wrong exceeded individual mean confidence *when correct* on the last quiz! Over $Time1_{(Quizzes1\text{-}5)}$ and $Time2_{(Quizzes7\text{-}11)}$, the increases in both the individual and group confidence when wrong are statistically significant (individual $t_{(31)} = -3.593$, $p = .001$; group $t_{(8)} = -2.606$, $p = .031$). Thus, even when participants receive individualized feedback each week, along with explicit instructions to lower their confidence scores when they are the slightest bit uncertain of their answers, students working as individuals or in their groups tend to become exceeding more (rather than less) confident in their wrong answers! Figure 5 demonstrates the relative mean confidence scores for both groups and individuals across the 11 quizzes. During these quizzes, participants in groups consistently demonstrate more confidence than participants working as individuals, especially when they are wrong ($t_{(10)} = 4.819$, $p = 0.001$)!

Insert Figure 5 about here.

The correlations between individual and group confidence when correct and wrong is likewise positive, rather than negative as we had hoped. That is, if the students had progressively increased their confidence when correct over the quizzes while decreasing their confidence when they suspected they might be wrong (the situation depicted in Figure 1), the correlations between confidence correct and confidence wrong would be negative, and increasingly so over the quizzes. Instead, individuals (and groups) giving high ratings of confidence when correct tend also to give high confidence ratings when wrong. (See Table 5.)

Insert Table 5 about here.

Probability of Correct or Wrong Answers in Groups of 3. Using groups of 3 members, we next examined the influence of group answer selection when group members enter their groups with individual item choices that are a combination of correct-correct-wrong (CCW) or correct-wrong-wrong (CWW). The top panel of Figure 6 separates group answers and associated confidences into their prior individual answer and confidence patterns for CCW groupings; the bottom panel does so for CWW groupings. The probability of a CCW group answering correctly is 0.94!

Insert Figure 6 about here.

During the 0.06 times that the group chooses an incorrect answer, the single group member with the wrong choice was more confident of that choice than the pair of group members who were previously correct. Thus, it appears that being highly confident and wrong influences the group answer selection more than moderate confidence when correct, even when coupled with 2 similar (correct) answers.

If 2 members have wrong answers and 1 member is correct, the CWW groups choose a correct answer slightly less than half of the time, as shown in the lower portion of Figure 6. Moreover, the mean confidence level of the CWW groups that choose correctly is similar to the CWW groups that choose a *wrong* answer, an outcome quite different from that obtained for the

CCW groups that choose correct and wrong answers, as shown in the top portion of Figure 6. It thus appears that even when CWW groups choose correctly, the effect of having 2 wrong answers dampens the confidence levels of the group.

We next examined the effects of similar versus dissimilar wrong answers in the CWW groups to see whether similar wrong answers suggest a majority effect that may influence group answer selection. Groups that have 3 dissimilar answers (2 wrong and 1 correct), have a .60 chance of answering correctly. However, having two individuals in the group who have given the same wrong answer decreases the likelihood that the group will answer correctly down to .36, which is near chance.

Thus, it appears that group answer selection is influenced both by prior individual answer confidence and by a majority effect of similar (correct or wrong) answers. In groups of 3, individuals with high confidence, regardless of the appropriateness of their answer, tend to influence their group's answer selection (i.e., as in Figure 6 for the CCW groups). When 2 members of a group have similar wrong answers, group confidence is higher than when 2 members of a group have multiple wrong answers (i.e., as in Figure 7). That is, reaching

---

Insert Figure 7 about here.

---

consensus from multiple, divergent opinions, decreases the groups' mean confidence scores both when the group decision is correct and when it is wrong.

Discussion

The main finding of this investigation is that even when a great deal of informative feedback is given to individuals and groups about their actual answer accuracy and understanding, overconfidence in wrong answers remains high. Most importantly, it appears that over time and the series of quizzes, groups tend to become more and more convinced of their own understanding of the information, not only when they are correct, but also when they are wrong! Additional analyses and future research need to be done to determine what other effects, if any, gender, student achievement in the course (i.e., upper and lower quartile rankings) and items requiring higher level reasoning strategies have on groups' assignment of confidence and decision making processes.

The effects of harder quizzes, familiarity of taking the same quiz twice, and maintaining the original groups throughout a course also need to be examined by further research. Likewise, discussion of answer rationales in the groups may have an effect on the assignment of group confidence scores. To examine whether discussing answer rationales has an effect on confidence assignment, individuals could write out their rationale for selecting each answer prior to marking their individual confidence score. Also, additional research is needed to investigate whether longer reflection times are involved in selection of answers for items on which participants report lower confidence scores. The length of time that participants need to answer questions and give confidence ratings both as individuals and as members of a group may be negatively correlated with their confidence ratings. Further examination of individual and group differences in confidence profiles, rationales, and accuracy may offer insights to the general problem of highly confident, but wrong.

To summarize, the present investigation combines the processes and research literatures of human judgment and metamemory (e.g., Lichtenstein & Fischhoff, 1977) with cooperative learning groups (e.g., Johnson, D. W. & Johnson, R., 1989). We hypothesized that placing students into groups would improve not only the accuracy of the students' answers but also the calibration of their confidence judgments. Within learning groups, students do improve the accuracy of their answers, but they do not appropriately adjust their confidence judgments. Worse yet, the benefits of students working in groups appear to come at the cost of an unwanted increase in students' confidence that group decisions are correct when in fact they are wrong. Thus, the primary scientific importance of our research concerns the startling tendency toward overconfidence in assessments of knowledge even under conditions of intensive feedback and assignment of students to learning groups. The educational significance of the research lies in what it tells the instructor about the difficulty of helping students to more appropriately discriminate between what they do understand and what they do not. Surprisingly, neither frequent feedback nor assignment to structured learning groups, irrespective of their other benefits, appear effective at improving students' calibration of confidence and, in particular cases, may even make it worse.

References

Adams, J. K. (1957). A confidence scale defined in terms of expected percentages. American Journal of Psychology, 79, 432-436.

Cohen, E. G. (1986). Designing groupwork. New York, NY: Teachers College Press.

Cooper, J., Prescott, S., Cook, L., Smith, L., Mueck, R., & Cuseo, J. (1990). Cooperative learning and college instruction. Long Beach, CA: California State University Institute for Teaching and Learning.

Fox, P. W., & LeCount, J. (1991, April). When more is less: Faculty misestimation of student learning. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1977). Knowing with certainty: The appropriateness of extreme confidence. Journal of Experimental Psychology: Human Perception and Performance, 3, 552-564.

Johnson, D. W. & Johnson, F. (1991). Joining together: Group theory and group skills (4th ed.). Englewood Cliffs, NJ: Prentice-Hall, Inc.

Johnson, D. W., Johnson, R., & Holubec, E. (1990). Circles of learning: Cooperation in the classroom (3rd ed.). Edina, MN: Interaction Book Company.

Johnson, D. W., & Johnson, R. (1989). Cooperation and competition theory and research. Edina, MN: Interaction Book Company.

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. Journal of Experimental Psychology: Human Learning and Memory, 6, 107-118.

Lichtenstein, S., & Fischhoff, B. (1977). Do those who know more also know more about how much they know? Organizational Behavior and Human Performance, 20, 159-183.

Shaughnessy, J. J. (1979). Confidence-judgment accuracy as a predictor of test performance. Journal of Research in Personality, 13, 505-514.

Zechmeister, E. B., Rusch, K. M., & Markell, K. A. (1983). Training college students to assess accurately what they know and don't know. (Report No. NIE-G-81-0093). Chicago, IL: Loyola University. (ERIC Document Reproduction Service No. ED 238 337)

Table 1

Quiz Mean Accuracy Scores of Students Working As Individuals and in Groups over 11 Quizzes

Variable	Mean	Std Dev
Individual Accuracy	0.83	0.038
Group Quiz Accuracy	0.89**	0.043

Note. Means differ significantly at $p < .001$.

Table 2

Mean Accuracy of Students Working As Individuals and in Groups over Quizzes 1-5 (Q1-5) and Quizzes 7-11 (Q7-11)

Variable	N	Mean	Std Dev
Individual Accuracy	Cases		
Time1 (Q1-5)	32	.832	.075
Time2 (Q7-11)	32	.833	.075
Group Quiz Accuracy	Quizzes		
Time1 (Q1-5)	5	.89	.036
Time2 (Q7-11)	5	.88	.056

Table 3a

Overall Mean Confidence Scores of Students Working As Individuals and in Groups during the 11 Quizzes

Variable	Mean	Std Dev
Individuals		
Overall Confidence	77.01	19.25
Confidence Correct	80.15	18.06
Confidence Wrong	65.22	19.02
Groups		
Overall Confidence	88.78	14.77
Confidence Correct	90.80	12.73
Confidence Wrong	71.90	18.04

Table 3b

Quiz Mean Confidence Scores of Students Working As Individuals and in Groups on the 11 Quizzes

Variable	Mean	Std Dev
Individuals		
Overall Confidence	77.00†	2.71
Confidence Correct	80.14‡	2.80
Confidence Wrong	65.03**	4.32
Groups		
Overall Confidence	88.65†	2.50
Confidence Correct	90.75‡	2.31
Confidence Wrong	70.27**	5.40

Note. Means with matching symbols differ significantly at $p \leq .001$.

Table 4

Mean Confidence Scores of Students Working As Individuals and in Groups over Time1 (Quizzes 1-5) and Time2 (Quizzes 7-11)

Variable	N	Mean	Std Dev
Individuals	Cases		
Overall Confidence			
Time1 (Q1-5)	32	75.30†	8.38
Time2 (Q7-11)	32	78.76†	8.50
Confidence Correct			
Time1 (Q1-5)	32	78.47	7.77
Time2 (Q7-11)	32	81.09	7.93
Confidence Wrong			
Time1 (Q1-5)	32	61.74‡	12.22
Time2 (Q7-11)	32	68.87‡	11.11
Group	Quizzes		
Overall Confidence			
Time1 (Q1-5)	5	87.59	2.39
Time2 (Q7-11)	5	89.85	2.56
Confidence Correct			
Time1 (Q1-5)	5	90.05	1.73
Time2 (Q7-11)	5	91.65	2.91
Confidence Wrong			
Time1 (Q1-5)	5	66.61**	1.95
Time2 (Q7-11)	5	73.93**	5.96

Note. Means with matching symbols differ significantly at $p < .05$.

Table 5

Correlations Between Confidence When Correct and Confidence When Wrong for Groups and Individuals at Time1 (Quizzes 1-5) and Time2 (Quizzes 7-11)

Variable	N	r
Individuals	Cases	
Cf Correct v. Wrong		
Time1 (Q1-5)	32	.75
Time2 (Q7-11)	32	.62
Group	Quizzes	
Cf Correct v. Wrong		
Time1 (Q1-5)	5	.56
Time2 (Q7-11)	5	.42

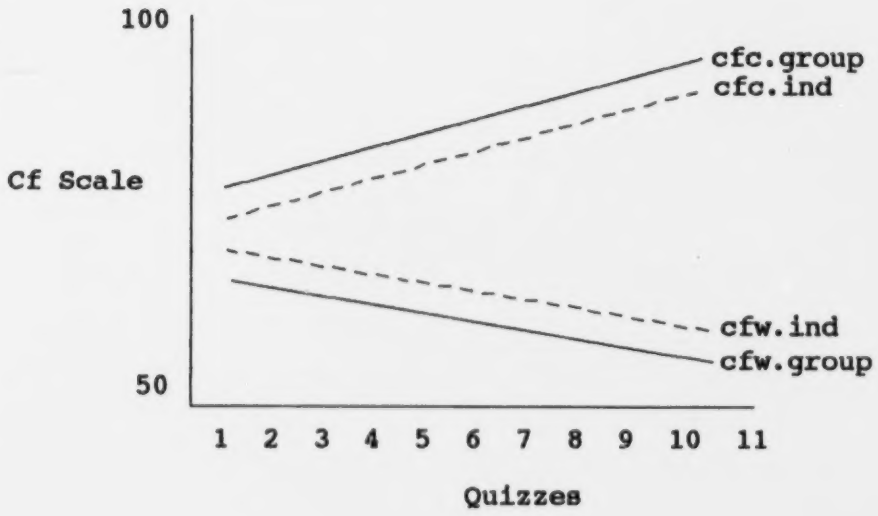


Figure 1. Predicted individual and group confidence when correct (cfc) and when wrong (cfw).

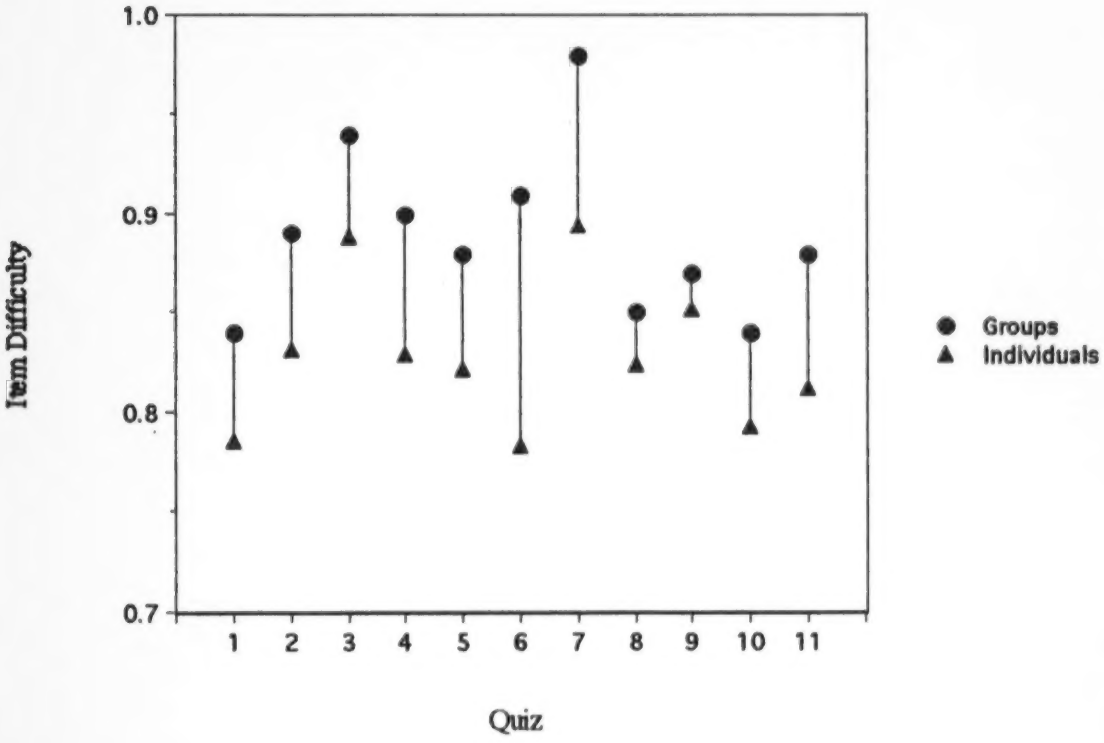


Figure 2. Mean accuracy scores for individuals and groups on Quizzes 1-11.

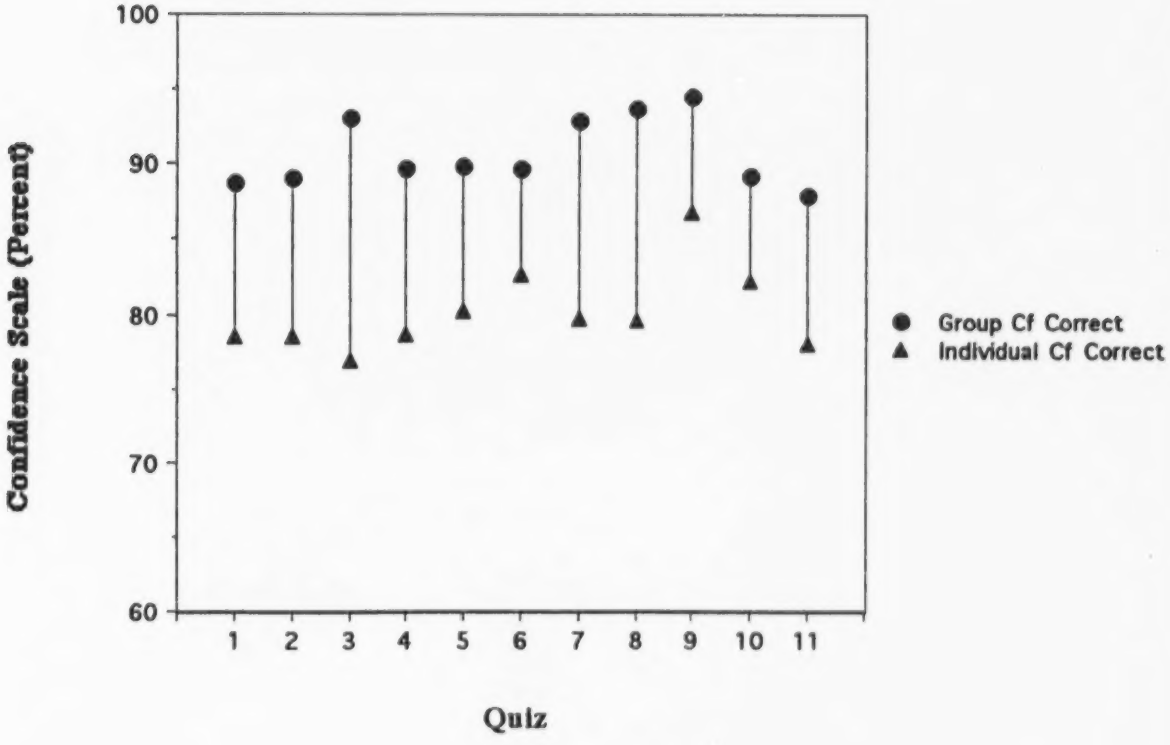


Figure 3. Mean confidence when correct ratings for individuals and groups on Quizzes 1-11.

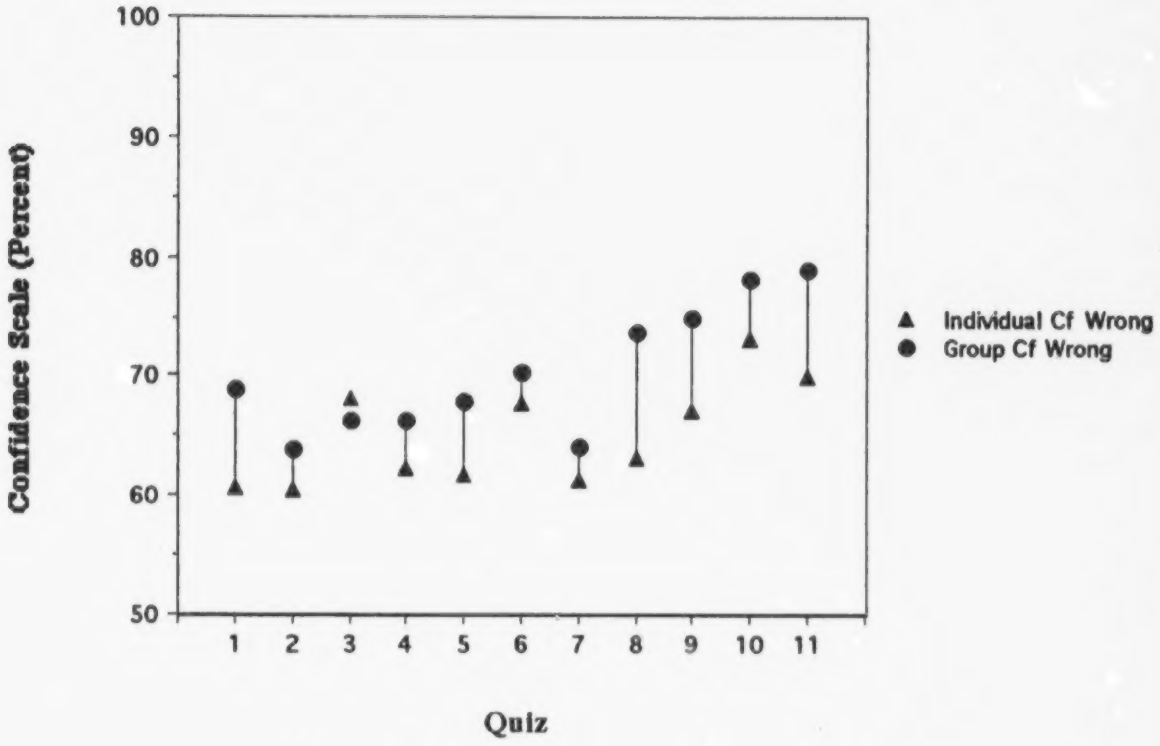


Figure 4. Mean confidence when wrong ratings for individuals and groups on Quizzes 1-11.

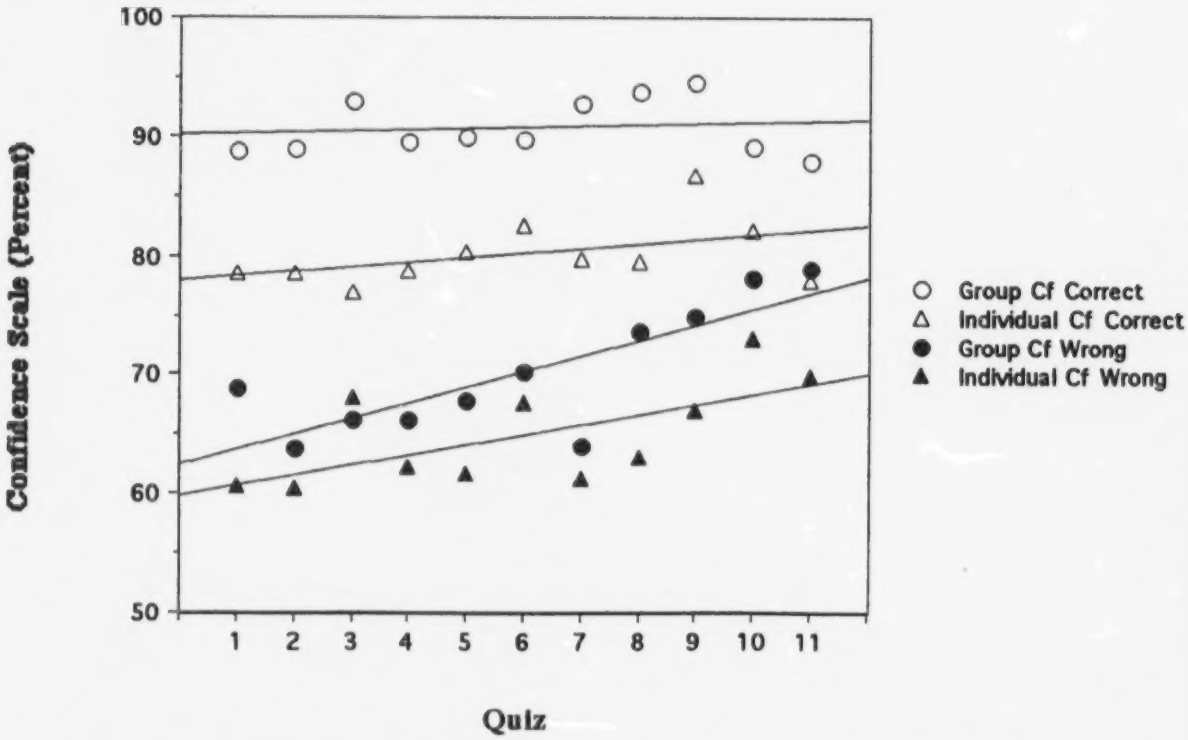


Figure 5. Individual and group mean confidence ratings when correct and wrong for individuals and groups on Quizzes 1-11.

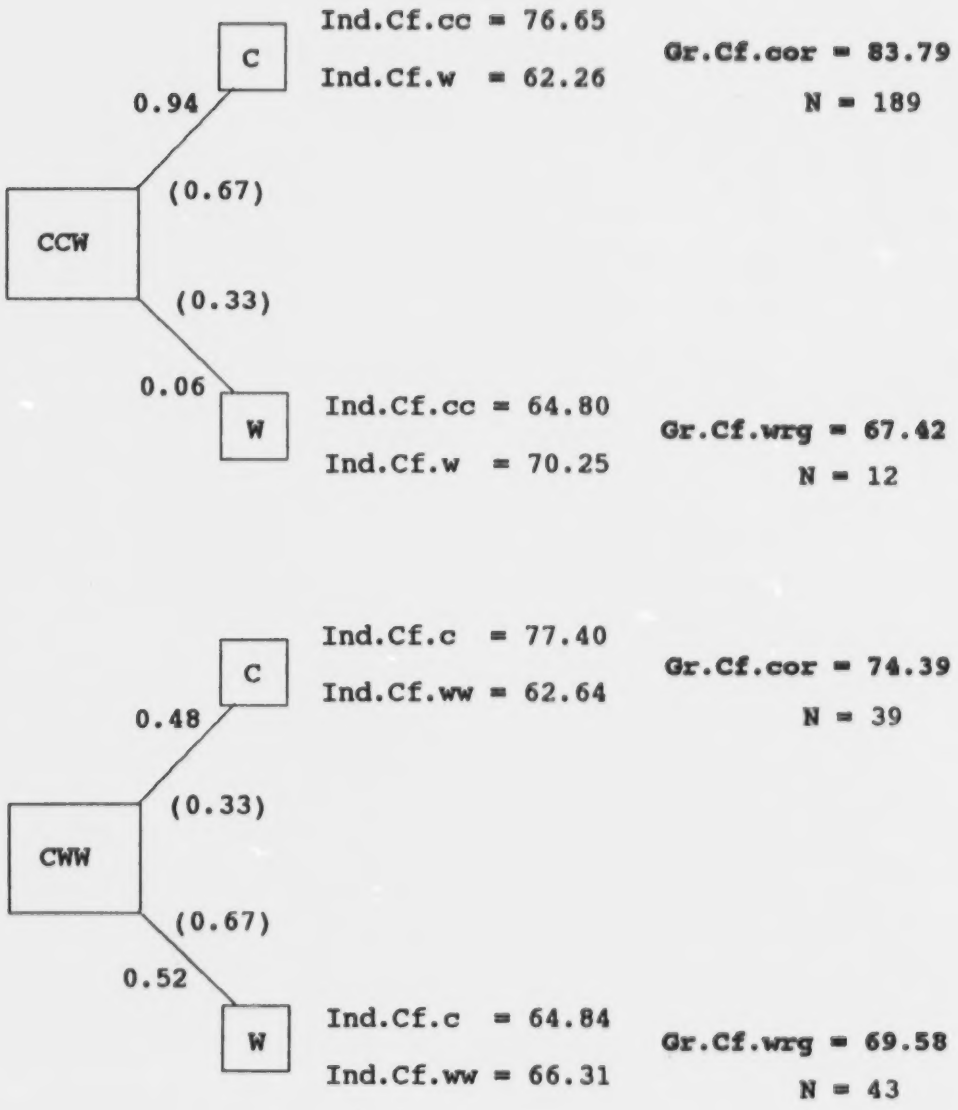
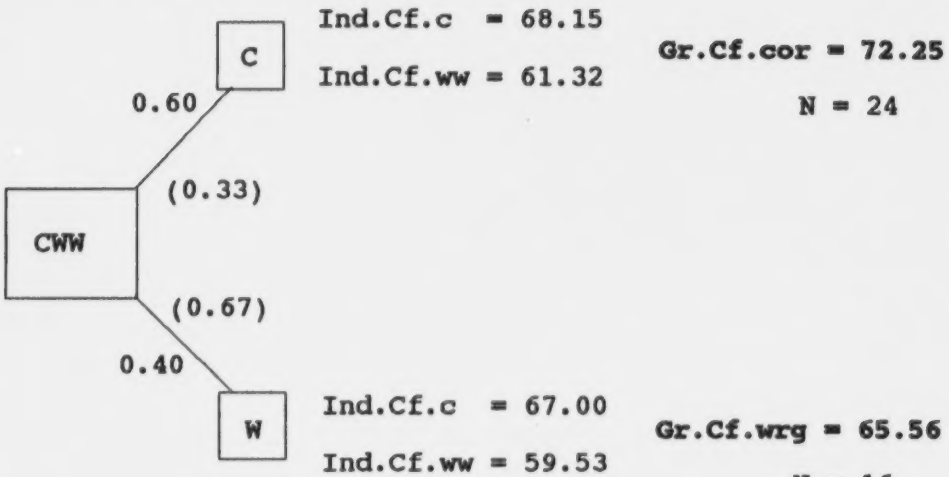


Figure 6. Accuracy probabilities and mean confidence scores for individuals and their groups of 3 when group members have dissimilar answers. Expected probabilities are in parentheses.

Dissimilar Wrong Answers



Similar Wrong Answers

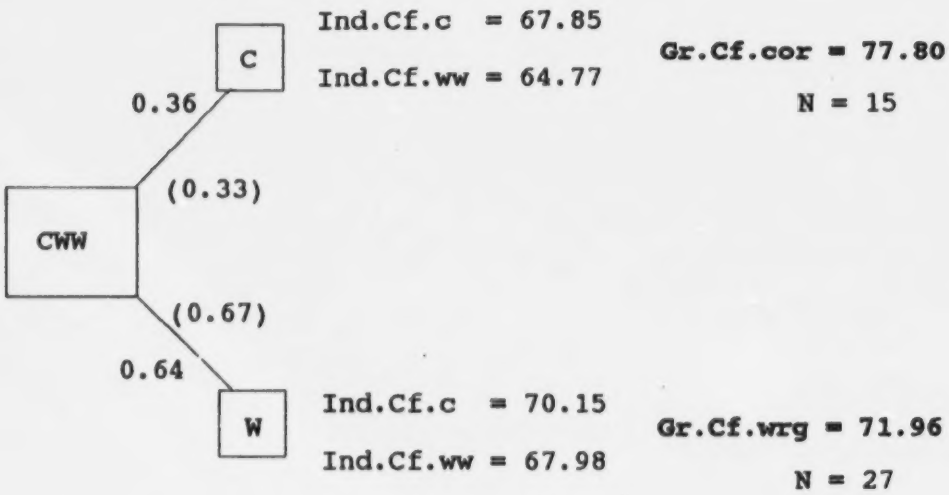


Figure 7. Accuracy probabilities and mean confidence scores for individuals and their groups of 3 when group members have similar and dissimilar answers. Expected probabilities are in parentheses.